

This listing of claims will replace all prior versions, and listings, of claims in the application.

**Listing of Claims:**

Claims 1-17. (Canceled).

18. (Previously Presented) A method of controlling a system to optimize an objective function thereof, the system performing a plurality of candidate actions and monitoring response performances of a performance of a respective candidate action, where the objective function is a function of the monitored response performances following decisions and actions taken, the method comprising the steps of:

a) monitoring response performance of a respective candidate action that is chosen to be performed by the system;

b) storing, according to the candidate action performed by the system, a representation of said monitored response performance;

c) calculating the expected growth in regret associated with each of the plurality of candidate actions, assessed using a probability distribution based on the historical response performances to date of said plurality of candidate actions, where the expected growth in regret is a system performance measure that is calculated to represent the trade-off between the relative merit of exploration of one or more apparently non-best candidate actions to mitigate the risk of ignoring one of said one or more apparently non-best candidate actions which may actually be the current best candidate action, with respect to the relative merit of exploiting what appears to be the current best candidate action but which in fact may not be the current best candidate action, based on said historical response performances to date;

d) choosing as the next action the candidate action that is calculated to result in the lowest expected growth in regret after the chosen candidate action is performed by the system;

e) commanding the system to perform the chosen next action; and

f) repeating steps a) to e) to control the system so as to substantially optimize the objective function of the system.

19. (Previously Presented) A method according to claim 18 wherein step c) includes assessing which candidate action is likely to result in the lowest expected growth in

regret on the basis of a true best candidate action which has the mean of said probability distribution.

20. (Previously presented) A method according to claim 18 wherein step c) includes evaluating the cost or losses associated with presenting a lower performing candidate action and the gain or benefit associated with knowing the true position of the current best observed candidate action on said probability distribution.

21. (Previously presented) A method according to claim 20 wherein step c) includes assessing which candidate action is likely to result in the lowest expected growth in regret according to an assumption that the current best observed candidate action is assumed to have zero uncertainty around its mean or expected response performance.

22. (Previously presented) A method according to claim 18 wherein step c) includes assessing which candidate action is likely to result in the lowest expected growth in regret according to an assumption of a Student's distribution and evaluation of Student's  $t$  parameters as the basis for estimating probabilities of unequal or equal response states between the candidate action with the current expected best response performance and any other candidate action.

23. (Previously Presented) A method according to claim 18 wherein step c) includes using a Monte Carlo algorithm to provide understanding of the probability distribution of the response performance of all of the plurality of candidate actions and either choosing the candidate action that if not taken would contribute most to an expected regret estimate, or choosing a candidate action with probability proportional to its contribution to the expected regret estimate if not taken.

24. (Previously Presented) A method according to claim 18 further comprising the step of:

g) applying a temporal depreciation factor to the stored representations of the response performance in order to depreciate the significance of the stored representations over time.

25. (Previously Presented) A method according to claim 24 wherein step g) includes applying, for each candidate action, a different temporal depreciation factor to the stored representations of the response performance thereof.

26. (Previously Presented) A method according to claim 18 further comprising the step of:

g) forcing the performance of each candidate action a minimum number of times or at a minimum rate.

27. (Canceled)

28. (Canceled)

29. (Canceled)

30. (Previously presented) A method according to claim 18 wherein the monitored response performance of a respective candidate action in step a) is stored in step b) in a form to enable sharing of the stored representation of said monitored response performance with another system.

31. (Previously Presented) A system having means for performing a plurality of candidate actions and means for monitoring response performances of a performance of a respective candidate action during performance of an objective function of the system, where the objective function is a function of the monitored response performances following decisions and actions taken, the system further having a control apparatus that is programmed to control the objective function of the system by performing the method comprising the steps of:

- a) monitoring response performance of a respective candidate action that is chosen to be performed by the system;
- b) storing, according to the candidate action performed by the system, a representation of said monitored response performance;
- c) calculating the expected growth in regret associated with each of the plurality of candidate actions, assessed using a probability distribution based on the historical response performances to date of said plurality of candidate actions, where the expected growth in regret is a system performance measure that is calculated to represent the trade-off between the relative merit of exploration of one or more apparently non-best candidate actions to mitigate the risk of ignoring one of said one or more apparently non-best candidate actions which may actually be the current best candidate action, with respect to the relative merit of exploiting what appears to be the current best candidate action but which in fact may not be the current best candidate action, based on said historical response performances to date;
- d) choosing as the next action the candidate action that is calculated to result in the lowest expected growth in regret after the chosen candidate action is performed;
- e) commanding the system to perform the chosen next action; and
- f) repeating steps a) to e) to control the system so as to substantially optimize the objective function of the system.

32. (Previously Presented) A robot comprising the system according to claim 31, where the control apparatus of the system controls the objective function of the robot so as to optimize the objective function of the robot.

33. (Currently Amended) A control apparatus for controlling a system to optimize an objective function thereof, the system of performing a plurality of candidate actions and monitoring response performances of a performance of a respective candidate action, where the objective function is a function of the monitored response performances following decisions and actions taken, the control apparatus comprising:

- a) means for monitoring response performance of a respective candidate action that is chosen to be performed by the system;

b) means for storing, according to the candidate action performed by the system, a representation of said monitored response performance;

c) means for calculating the expected growth in regret associated with each of the plurality of candidate actions, assessed using a probability distribution based on the historical response performances to date of said plurality of candidate actions, where the expected growth in regret is a system performance measure that is calculated to represent the trade-off between the relative merit of exploration of one or more apparently non-best candidate actions to mitigate the risk of ignoring one of said one or more apparently non-best candidate actions which may actually be the current best candidate action, with respect to the relative merit of exploiting what appears to be the current best candidate action but which in fact may not be the current best candidate action, based on said historical response performances to date; ~~and~~

d) means for choosing as the next action the candidate actions that is calculated to result in the lowest expected growth in regret after the chosen candidate action is performed by the system; and

e) means for commanding the system to perform the chosen next-action,  
wherein the control apparatus controls the system so as to substantially optimize the objective function of the system.

34. (Previously Presented) A method according to claim 18 wherein the representation of said monitored response performance contains at least one variable that characterizes the conditions under which the candidate action was performed.

35. (Previously Presented) A method of controlling a system with two or more subsystems to optimize an objective function of the system, the system being capable of performing a plurality of candidate actions, wherein a candidate action is represented by the selection of a lower level subsystem from said two or more subsystems, and wherein the system is capable of monitoring the response performance of the selected subsystem, where the objective function is a function of the monitored response performances following decisions and actions taken, the method comprising the steps of:

- a) monitoring response performance of a respective candidate action that is chosen to be performed by the system;
- b) storing, according to the candidate action performed by the system, a representation of said monitored subsystem performance in response to the candidate action;
- c) calculating the expected growth in regret associated with each of the plurality of candidate actions, assessed using a probability distribution based on the historical response performances to date of said plurality of candidate actions, where the expected growth in regret is a system performance measure that is calculated to represent the trade-off between the relative merit of exploration of one or more apparently non-best candidate actions to mitigate the risk of ignoring one of said one or more apparently non-best candidate actions which may actually be the current best candidate action, with respect to the relative merit of exploiting what appears to be the current best candidate action but which in fact may not be the current best candidate action, based on said historical response performances to date;
- d) choosing as the next action the candidate action that is calculated to result in the lowest expected growth in regret after the chosen candidate action is performed by the system;
- e) commanding the system to perform the chosen next action using a corresponding lower level subsystem; and
- f) repeating steps a) to e) to control the system so as to substantially optimize the objective function of the system.